Maturitas 79 (2014) 401–412

Contents lists available at ScienceDirect

Maturitas

journal homepage: www.elsevier.com/locate/maturitas

Review

A systematic review of the longitudinal relationships between subjective sleep disturbance and menopausal stage

Qunyan Xu[a,*], Cathryne P. Lang[b], Naomi Rooney[c]

[a] Sansom Institute of Health Research, School of Nursing and Midwifery, University of South Australia, GPO Box 2471, Adelaide 5001, SA, Australia
[b] School of Psychology, Australian Catholic University, PO Box 456, Virginia, QLD 4014, Australia
[c] School of Nursing and Midwifery, University of South Australia, GPO Box 2471, Adelaide 5001, SA, Australia

ARTICLE INFO

Article history:
Received 12 September 2014
Accepted 20 September 2014

Keywords:
Menopausal stage
Sleep disturbance
Longitudinal
Systematic review

ABSTRACT

Sleep disturbance is a common complaint for women going through the menopausal transition. A previous systematic review and meta-analysis of cross-sectional studies showed a small but significant relationship between self-reported sleep disturbance and menopausal stage and highlighted a possible influence of culture. However, the longitudinal relationship between self-reported sleep disturbance and menopausal transition has not been explored. This paper aimed to review literature on the longitudinal relationship between self-reported sleep disturbance and menopausal transition among community dwelling midlife women. Multiple electronic databases were systematically searched. Literature published prior to 2013 was reviewed. A narrative synthesis was used to analyse the results due to high level of heterogeneity across the included studies. Overall, review of eligible studies showed a small increased risk of self-reported sleep disturbance as women go through the menopausal transition after adjustment of potential confounders. Although the methodological quality of the majority of included studies was classified as high, the impact of culture on this relationship could not be explored, as all of the included studies were conducted in western countries. Like vasomotor symptoms, self-reported sleep disturbance is one of the core menopausal symptoms. Management strategies should be put in place to help women manage sleep disturbance to prevent complications and to improve health related quality of life.

Contents

* Corresponding author. Tel.: +61 8 8302 2531; fax: +61 08 8302 2168.
E-mail addresses: andie.xu@unisa.edu.au (Q. Xu), cathryne.lang@acu.edu.au (C.P. Lang), naomi.margetts@unisa.edu.au (N. Rooney).

http://dx.doi.org/10.1016/j.maturitas.2014.09.011
0378-5122/

1. Introduction

Sleep disturbance is a common health concern, with 30% of the general population experiencing sleep difficulty of varying severity [1]. While sleep disturbance is more common for women than men at any stage of their life [e.g., 1, 2, 3], during the menopausal transition, the prevalence of sleep disturbance increases dramatically from 30% in premenopausal women to approximately 50% in peri- and postmenopausal women [for reviews, see 3, 4].

At a physiological level, the decreasing levels of Estradiol and increasing follicle stimulating hormone (FSH) levels may interfere with the secretion of melatonin and other circadian hormones [3]. In addition, ageing may be an important consideration. As women age, their circadian rhythm changes, resulting in shortened sleep duration and early morning awakening.

However, it is clear that sleep disturbance is multifactorial, and many aspects of the menopausal transition are believed to have an association with sleep disturbance in addition to the hormonal changes of the menopausal stages [3–6]. First, vasomotor symptoms, particularly night sweats, are associated with difficulty maintaining sleep at night. The sudden perspiration can awaken women, as well as necessitate the change of bed linen and clothes. Furthermore, women's psychosocial status is also a possible contributor to sleep disturbance, as sleep disturbance is often an inherent feature of mental health problems; some of which are also more common in women. Given the adjustments many women have to make during menopausal transition, psychosocial status should always be considered. Lastly, health behaviour such as weight status, physical activity, consumption of alcohol and tobacco can influence women's sleep as well. A high body mass index is associated with obstructive sleep apnea, which is a cause for sleep disturbance. Individuals who complete moderate physical activity during the day experience better sleep quality. Given the associations between these factors and sleep disturbance, it is essential that when investigating the relationship between menopausal transition and sleep disturbance, these factors is taken into consideration.

The question of whether menopausal stage independently predicts sleep quality was examined via systematic review and meta-analysis in our earlier paper [7]. This paper quantified the relationship between menopausal transition and sleep disturbance, and more importantly, highlighted the cultural variation in this relationship. However, our conclusions are based on the analysis of cross-sectional data, which is does not allow the consideration of the influence of sleep disturbance at pre-menopausal stage on sleep at a later time, or changes of confounding factors (such as vasomotor symptoms, psychosocial factors or health variables) over time. It may be important to control for women's sleep quality before their menopausal transition because women who sleep poorly may be more likely to experience sleep disturbance at menopausal transition than those who do not. Investigating the changes in confounders allows for the evaluation of whether the increase in sleep disturbance during the menopausal transition is a result of changing life circumstances, health behaviour or other factors.

The purpose of this paper is to systematically review the longitudinal relationship between sleep disturbance and the menopausal transition among community dwelling women using a narrative synthesis approach. Critique of the methodological quality of eligible studies will be conducted to inform the interpretation of the overall findings.

2. Method

2.1. Literature search

A systematic literature search was conducted in multiple electronic databases including AgeLine, CINAHL, Cochrane Library, Health Collection, Health and Medical Complete, Informit Health Collection-Health Collection, Joanna Briggs Institute, PubMed, ProQuest Psychology Journals, PsycARTICLES, Psychology and Behavioral Sciences Collection, PsycINFO, ProQuest Family Health, and SAGE text. *Sleep* or *sleep disorder* in combination with *menopause* or *menopausal* were used as both keywords and Medical Subject Headings/subject terms. The search was conducted from the first available year to November 2013. The search was limited to human subjects, apart from this; no other limits were applied in search strategy (including language) to maximize search outcomes. The reference lists of review articles were manually searched to ensure eligible studies were included.

2.2. Inclusion and exclusion criteria

The purpose of this review was to evaluate the longitudinal relationship between menopausal transitions and sleep disturbance in community dwelling midlife women. Studies were included if they (1) quantitatively examined the longitudinal relationship between menopausal transition and sleep disturbance; and (2) the comparison of sleep disturbance was performed using the premenopausal (reproductive) stage or a combined premenopausal and early perimenopausal stage as the reference group.

Studies that met any of the following criteria were excluded: (1) failed to define stages of menopausal transition clearly; (2) examined women at one menopausal stage only (e.g. postmenopausal women only); (3) classified menopausal stages in a way that is not comparable with other included studies (e.g. non-postmenopausal vs. postmenopausal); (4) examined a specific sleep disorder such as restless leg syndrome, nocturia or obstructive sleep apnea; (5) examined the result of sleep disturbance such as daytime sleepiness rather than sleep disturbance itself; and (6) used a patient sample. Papers which did not present primary research containing quantitative data, such as traditional reviews, commentaries or editorials were also excluded. When multiple reports were published from a single study (defined as one study sample), only the report that contains relevant and the most complete data was included. Fig. 1 provides a flowchart displaying the paper selection process.

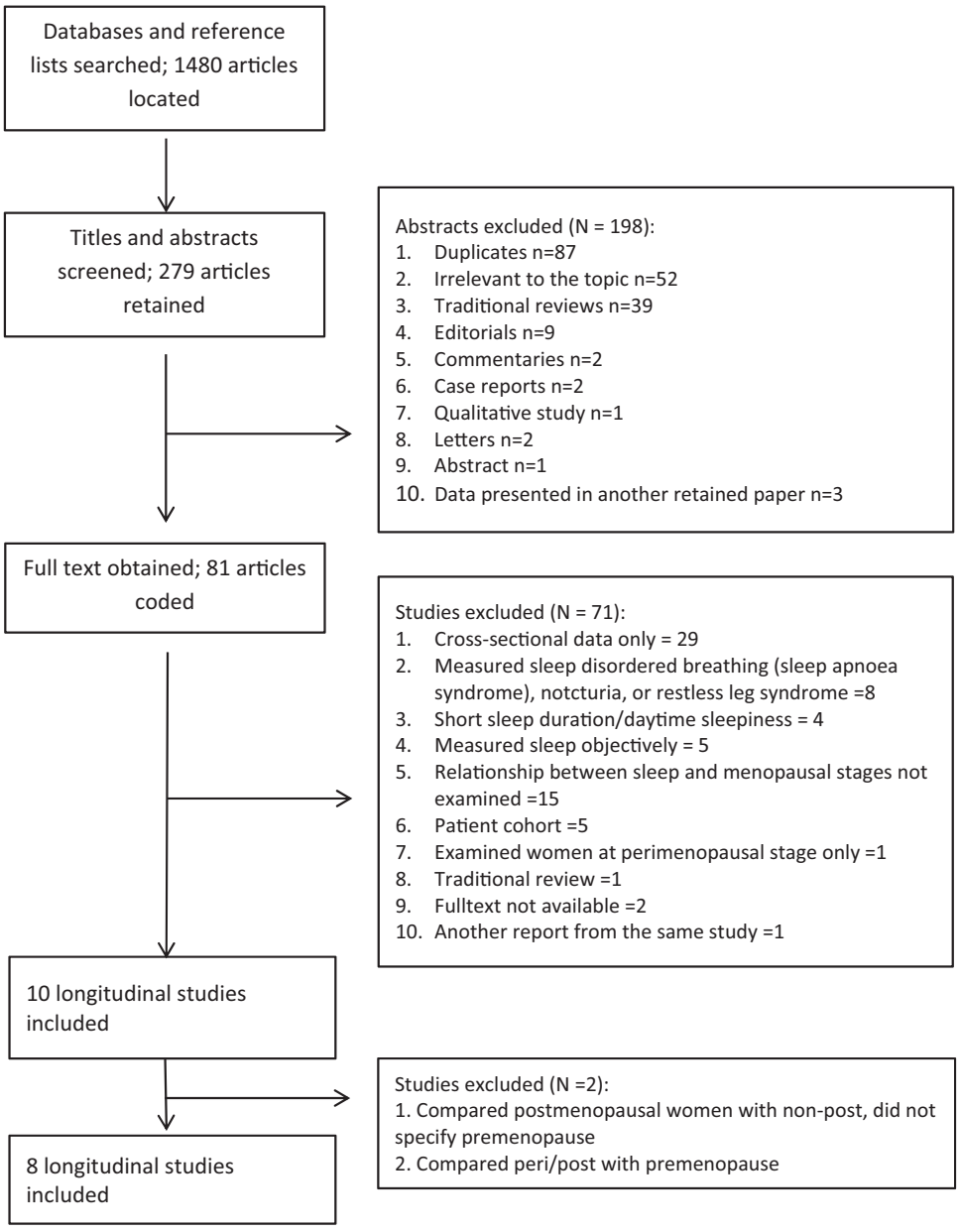

Fig. 1. Flow chart of search, retrieval and selection process.

2.3. Description of study characteristics

For each of the included studies, the following characteristics were obtained: ethnicity of the sample, inclusion and exclusion criteria, sampling strategy and representativeness of study sample of the target population, sample sizes at both baseline and final follow up point, duration and frequency of follow up for each of the key variables—sleep disturbance, menopausal transition and confounders, measurement of sleep disturbance and menopausal stages, adjustment of confounders and outcome of the relationship between sleep disturbance and menopausal stages.

2.4. Assessment of risk of bias in included studies

Risk of bias in included studies was assessed according to the Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 in relation to non-randomized studies [8]. Assessment for risk of bias was conducting for the following types; selection bias (comparability of baseline sample characteristics between menopausal stages and response rate), reporting bias (selective reporting of significant results), and information bias related to measurement of key variables, attrition bias and adjustment for confounders. Risk of performance bias and detection bias were not evaluated as they are irrelevant to studies without intervention. Risk of bias assessment was undertaken by two authors independently (Xu & Lang), and disagreement was discussed and solved.

2.5. Data synthesis

Type and magnitude of the effect size of the relationship between menopausal transition and sleep disturbance was extracted. Meta-analysis of the data was not performed, as there was a significant variation effect sizes presented as well as an

overall small number of studies. Included studies were analysed using narrative synthesis.

3. Results

3.1. Characteristics of included studies

These longitudinal data were collected in Australia [9,10], the United States of America [11–15], and the United Kingdom [16]. Three papers were published in the last 5 years [9,14,16], and another three papers in the prior 5-year period between 2004 and 2008 [10,11,13]. A total of 13025 participants from these eight studies were included for analysis. The majority of the studies' samples were Caucasian women, although two studies targeted African American women [11,13] and the Study of Women Across the Nation (SWAN) [11] also targeted other ethnic minorities (recruiting 229 Hispanic, 244 Chinese American, and 275 Japanese American women).

Most studies used postal survey as the primary method of data collection. A trained interviewer (e.g. nurse) was involved for the purpose of screening eligible study participants [14], or in obtaining biological sample (e.g. blood) when necessary [13].

3.2. Measurement of sleep disturbance

Out of the eight included studies, five measured sleep disturbance as a global construct [9,13–16]. Four of these studies used a single item assessing sleep disturbance over varying periods of time [9,10,12,16] (Table 1). The Penn Ovarian Aging Study [13] used a factor score derived from the St Mary's Sleep Questionnaire, named 'sleep quality'. This sleep quality score was chosen over other factors because it accounted for 37% of the variance of the entire questionnaire. Three studies measured common sleep problems that are usually associated with insomnia such as difficulty initiating sleep (DIS), difficulty maintaining sleep (DMA) and early morning awakening (EMA) separately [10–12].

Assessment of sleep disturbance has been focused on the frequency of occurrence [9,11,12,15], severity/bothersomeness of sleep disturbance symptoms [14,16], or a combination of both aspects [10,13]. The recall time frame over which sleep disturbance was assessed ranged widely from the previous night [13,15] to the last 12 months [9,16]. Two included studies did not state over what time frame sleep disturbance was assessed [10,14].

3.3. Measurement of menopausal stages

Stages of menopausal transition were classified using the STRAW criteria in most [9–12,14–16], but not all, cases. While the Penn Ovarian Aging Study [13] was described as following the STRAW criteria guidelines, there were some differences between the categories used in this study and STRAW criteria. For example, the criterion of '1 observed change in cycle length' does not meet the STRAW criterion of 'persistent' changes in cycle length; and there is some overlap between the early transition criterion of '≥60 days amenorrhea' and the late transition criterion of 'greater than or equal to 3 months of amenorrhea' (Table 3).

Despite classifying the stages of the menopausal transition according to STRAW criteria, a number of different terms were used to describe each stage. The most obvious difference was that in one study, premenopause was referred as the reproductive phase [10]. Consistent with STRAW, other studies broke perimenopause into early and late transition [10,13,14] whereas other studies grouped early and late transition into a single perimenopause group [9,11,12,15,16].

'Menopausal status not otherwise specified' as in the STRAW statement was also investigated with regard to its relationship with sleep disturbance. These menopausal statuses included surgical menopause, hormone replacement therapy (HRT) users and oral contraceptive pill users. In terms of natural vs. surgical postmenopause, two studies treated these as separate groups [11,16], whereas the Wisconsin Sleep Cohort Study [15] combined these two groups into one. Surgical menopause was not examined in the other five included studies [9,10,12–14]. In terms of HRT users, all of the studies treated women on HRT as a separate group except for the Penn Ovarian Aging Study which failed to report the analysis of HRT users [13] and the Australian Longitudinal Study of Women's Health [9] that excluded HRT users from analyses.

3.4. Follow-up periods

The duration of follow up time and the number of follow up points were highly variable. The study by Owens and Matthews had only one follow up point, which was approximately 46 months post baseline [12]. The Australian Longitudinal Study of Women's Health [9] had two follow-up points which were about two years apart over 5 years, and the Wisconsin Sleep Cohort Study [15] also had two follow up points but were four years apart. Three studies included between approximately 6 and 8 annual follow-up points [10,11,16]. The Penn Ovarian Aging Study [13] collected data at approximately every 9 months in the first 5 years and annually in the last 4 years over a period of 8 years. On the other hand, the Seattle Midlife Women's Study first collected data 8 to 12 times annually, then quarterly each year [14]. The Australian Longitudinal Study of Women's Health study [9], Seattle Midlife Women's Study [14] and the Medical Research Council National Survey of Health and Development study [16] had additional follow-up time waves which did not include measures of sleep, and are therefore, excluded for the purposes of this review.

3.5. Main findings: Influence of menopause on sleep quality

Seven of the eight studies compared premenopause with perimenopause, and all eight studies compared premenopause with postmenopause. With regard to the ***transition to perimenopause***, three of the four studies that measured global sleep disturbance reported an increased odds (OR: 1.40–1.97) of experiencing sleep disturbance [9,10,16]. Two of the three studies examined common sleep problems found an increased odds of experiencing DIS (OR: 1.48–1.96) [11,15] and DMS (OR: 1.48–1.96) [11,14], and only the SWAN study [11] found an increased odds for EMA. When ***postmenopause was compared*** with premenopause, four of the five studies that measured global sleep disturbance reported increased odds of having sleep disturbance (OR: 1.40–2.33) [9,10,12,16]. Of the three common sleep problems, only DMS seemed to have a reliable significant association with the transition to postmenopause with two studies reporting significantly worse sleep [11,14]. For DIS, the Seattle Midlife Women's Health Study [14] reported better sleep at postmenopause whereas the SWAN study [11] found worse sleep. (The Wisconsin Sleep Cohort Study [15] did not find a difference in sleep quality). No differences were found for EMA by any of the studies.

Although the STRAW criteria do not comment on surgical menopause, two of the included studies grouped women with ***surgical menopause*** separately from natural postmenopause, and compared these women with premenopausal women [11,16]. For the SWAN study, both DIS and DMS were significantly more common in women with surgical menopause; although no difference

[1] From: Woods & Mitchell (1997) Pathways to depressed mood for midlife women: Observations from the Seattle Midlife Women's Health Study. Research in Nursing & Health, 20, 119–129

was found for EMA [11]. For the Medical Health Research Council National Survey of Health and Development, a greater number of women with surgical menopause had moderate and severe sleep problems [16] (Table 3).

3.6. Potential confounders

Existing literature has suggested a number of factors that can potentially moderate the relationship between sleep disturbance and menopausal transition. These associated factors could be grouped into broad categories of socio-demographic factors, health behaviours (e.g. physical activity, consumption of alcohol and caffeine), physical health (e.g. number of physical conditions diagnosed), psychological health (e.g. depression and anxiety), vasomotor symptoms (e.g. night sweat or hot flush), and biological hormones (e.g. estradiol, FSH) (Table 1). Among the six studies that controlled for confounders [9,11,13–16], all of them considered socio-demographic factors, psychological status and health behaviour. Physical health was controlled for in the SWAN study [11], Wisconsin Sleep Cohort Study [15], Seattle Midlife Women's Health Study [14], and the Australian Longitudinal Study of Women's Health [9]. Biological hormones, typically estradiol and FSH, were adjusted for in the Seattle Women's Health Study [14] and the Penn Ovarian Ageing Study [13], but were treated as another indicator of menopausal transition in the SWAN study [11]. Vasomotor symptoms such as night sweats and hot flushes were adjusted for in the Penn Ovarian Ageing Study [13], the Seattle Women's Health Study [14], the Australian Longitudinal Study of Women's Health [9] and the Medical Research Council National Survey of Health and Development [16]. Similar to the biological hormone levels, the frequency of vasomotor symptoms was also used as an indicator of menopausal stages in the SWAN study [11]. Lifetime history of sleep disturbance was only assessed in the Medical Research Council National Survey of Health and Development in the report by Tom and her colleagues [16]. Out of the aforementioned confounders, socio-demographic factors including education, marital status and parity, history of sleep disturbance were only assessed once at baseline and were treated as time-independent variables in the subsequent analysis. The remaining confounders including psychological status, physical health, hormone levels and vasomotor symptoms were assessed concurrently with sleep disturbance over time and analysed as time-dependent variables. Health behaviour has been analysed as both time-independent [11,16] and time-dependent (time varying) variable [9,13–15] depending on the study.

Consideration of confounders was unclearly articulated in two studies. In the study by Owens and Matthews [12], it was unclear if blood pressure and psychological health were controlled for in calculating the OR of sleep disturbance for postmenopausal women compared to premenopausal women. The Melbourne Women's Midlife Health Project [10] used structural equation modelling, where sleep disturbance was treated as both outcome and predicting variables, thus the confounders adjusted for in this particular regression analysis could not be identified.

3.7. Risk of bias in included studies

The key features of risk of bias in included studies are presented in Fig. 2. Studies meeting our inclusion and exclusion criteria have reasonable methodological quality.

3.7.1. Selection bias and baseline response rate

Risk of selection bias in all of the included studies was considered to be low, as in longitudinal studies, women's baseline sleep disturbance were treated as controls for that woman at peri and postmenopausal stages, which eliminates the effect of between

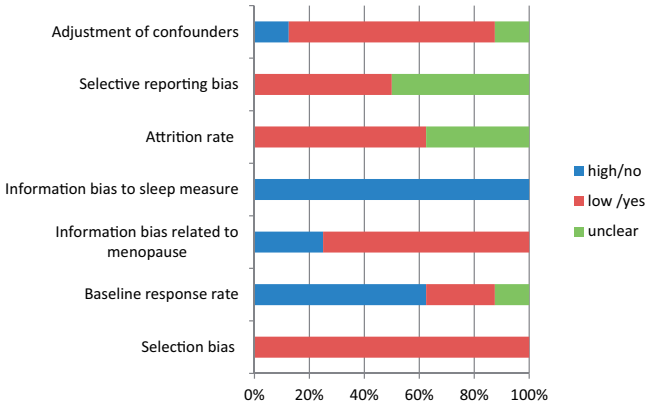

Fig. 2. Methodological quality graph: review authors' judgements about each methodological quality item presented as percentages across all included studies.

subject variation on the relationship between sleep disturbance and stages of menopausal transition. The baseline response rate was greater than 60% in two studies [13,16], approaching 60% in three studies [9,10,15], lower than 60% in the Seattle Midlife Women's Health Study [14], and not reported in another two studies [11,12].

3.7.2. Information bias related to menopausal status and sleep measures

Risk of information bias in relation to the assessment of menopausal stages was evaluated from three aspects—the consistency of the classification of menopausal stages with the STRAW criteria, the management of surgical menopause, and the management of HRT users. Studies that used a classification approach that is (1) consistent with the STRAW criteria, (2) excluded surgical menopause or analysed these women as a separate group, and (3) excluded HRT users or analysed these women as a separate group were considered to be at low risk of information bias. Studies that did not meet one or more of these three criteria were considered as having a high risk of information bias. According to this criteria, two studies were deemed to be at high risk of information bias related to measurement of stages of menopausal transition [13,15] and the remaining six studies were at low risk [9–12,14,16].

The risk of information bias in relation to sleep disturbance was considered high in all of the included studies, as sleep disturbance was measured on the basis of self-reported responses.

3.7.3. Attrition bias

Among studies where rates of attrition was reported or could be calculated, the rates of attrition ranged from 15.4% in a report of the Australian Longitudinal Study of Women's Health [9] to 27.2% in the SWAN study [11]. No relationship was found between duration of follow up and rates of attrition. Rates of attrition could not be determined in three studies [12,16]. No consensus exists with regard to what is an acceptable rate of attrition in longitudinal studies, as acceptable rates of attrition are highly dependent on the required sample size that allows detection of reliable research outcomes. Given that the sample sizes for analysis in the majority included studies were over 300, an arbitrary cut-off attrition rate of 30% was used to categorise the risk of attrition bias. According to this criterion, all of the six studies (see Table 2) with a clear reported attrition rate were deemed at low risk of attrition bias.

3.7.4. Selective reporting bias

The risk of selective reporting bias is rated as low if studies have an established public website for dissemination of research outcomes or a published study protocol. Four studies were rated as having low risk of selective reporting bias [9,11,14,16], and the

Table 1
Characteristics of the eight studies selected.

Author & year & study name	Ethnicity & country	Eligibility	Sampling & sample size	Measure of sleep disturbance	Confounders adjusted for	Analysis used	Effect size (95% CI)
Berecki-Gisolf et al. (2009) Australian Longitudinal Study of Women's Health (ALSWH)	Multiple, with the majority being Caucasian Australia	• Aged between 45–50 yrs • Agree to participate •Unknown menopausal status excluded	Randomly selected from the Medicare Database Data collected in 1996, 1998 & 2001 Baseline N (1996) = 8,612 End of F/U N (2001) = 7290	Composite measure Single item question: "*Have you experienced difficulty sleeping in last 12 months*" Response: "*never, rarely, sometimes, often*" 'sometimes' and 'often' coded as present	**Demographics** •Age •Age at menopause •Education •Financial stress **Psychological** •Clinically diagnosed depression •Mental health inventory •Life events **Physical health** •Self-rated overall health **Health behaviours** •BMI •Smoking **Menopausal symptoms** •Night sweats	Logistic regression model for repeated measures Logit link function and exchangeable correlation structure	Premenopause as the reference group Type of effect size: OR Peri (4 > y before menopause): 1.53 (1.32–1.77) Peri (1–4 y before menopause: 1.32 (1.18–1.46) Menopause (1 y before & after) 1.36 (1.20–1.55) Post (1–4 y after menopause)1.42 (1.22–1.66) Post (4–7 y after menopause) 1.51 (1.21–1.90) Post (>7 y after menopause) 1.38 (0.97–1.94)
Kravitz et al. (2008) Study of Women's health Across the Nation (SWAN)	Multi ethnics with the majority being Caucasian America, 7 community sites	• Aged between 42–52 yrs • Intact uterus and at least one intact ovary • Not currently on exogenous hormone • At least one menstrual period in the previous 3 months • Known race/ethnic group • Not pregnant	Varying recruitment strategies used at the 7 study sites. Data collected annually for 7 years Baseline *N* = 3045 End of F/U *N* = 2217	Individual sleep problems-DIS, DMS & EMA Question for each sleep problem "*How often do you experience sleep problems in the last 2 weeks?*" Response: "*none, < once/wk, 1-2/wk; 3-4/wk; > =5/wk*" "3 times/week and greater" coded as present	**Demographics** • Race • Age • Marital • Education • Income • Study site **Physical health** • Self-rated overall health • Number of medical conditions • Number of physical symptoms • Night time urinary • Pain score • BMI • Use of HT/psychotropic meds **Psychological status** • CESD • Anxiety • Social support • Perceived stress • Stressful life events **Health behaviours** • Smoking • Alcohol/caffeine use • Exercise	Random effects logistic regression	Premenopause as the reference group Type of effect size: OR ***DIS*** Early peri: 1.16 (0.93–1.45) Late peri: 1.49 (1.08-2.06) Post: 1.77 (1.29–2.44) Surgical: 2.05 (1.21–3.49) Post + HRT: 0.95 (0.57–1.57) Surgical + HRT: 1.05(0.58–1.88) ***DMS*** Early peri: 1.26 (1.08–1.47) Late peri: 1.96 (1.56–2.46) Post: 1.78 (1.42–2.23) Surgical: 2.04 (1.38-3.03) Post + HRT: 1.22 (0.97–1.54) Surgical + HRT: 1.13(0.82–1.57) ***EMA*** Early peri: 1.06 (0.88–1.28); Late peri: 1.53 (1.17–2.01) Post: 1.19 (0.91–1.56) Surgical: 1.50 (0.96–2.33) Post + HRT: 1.03 (0.70–1.53) Surgical + HRT: 1.60 (0.98–2.61)

Table 1 (*Continued*)

Author & year & study name	Ethnicity & country	Eligibility	Sampling & sample size	Measure of sleep disturbance	Confounders adjusted for	Analysis used	Effect size (95% CI)
Owens and Matthews (1998) No study name	Not reported USA	• Aged between 45–53 yrs • Premenopausal status • Not surgical menopause • Not peri or post	Convenience sampling of women from the Graduate School of Public Health and UP Sleep data collected at baseline & 46 mths post baseline Only reported 213 women for total analysis	Composite measure Single item question asking *if they had any trouble sleeping in the past 6 mths* Responses: *yes; no* Note: Individual sleep problems were asked, but data not presented	• Use of HRT at the follow up point • Hot flashes at follow up point interaction of the above two items	Logistic regression	Premenopause as the reference group Type of effect size: OR Peri: not reported Post: 2.23 ($p = 0.01$)
Pien et al (2008) Penn Ovarian Aging Study	Multi ethnicity (50% Caucasians & 50% African American) USA	• With menstrual cycles • Aged between 35–47 yrs • English speaker • Intact uterus and at least 1 ovary. • Not on psychotropic • Not on HRT • Not pregnant • No serious health issues • Not lactation	Random selected from household numbers in Philadelphia Data collected every 9mth in the first 5 yrs and annually in the later 4 yrs Baseline $N = 433$ End of F/U $N = 311$	Composite measure A factor called "sleep quality score" was used. This factor was derived from the St. Mary's Hospital Sleep Questionnaire (SMHSQ)	**Demographics** • Race • Age • Employment **Health behaviours** • Caffeine • Alcohol use • Smoking **Psychological status** • Zung anxiety • History of depression • CES-D scale • PSS • PMS **Menopausal symptoms** • Hot flashes **Hormones** • Estradial • Testosterone • FSH • Inhibin B & time	Linear mixed regression model Equicorrelated working structure	Premenopause as the reference group Type of effect size: sleep quality factor score (p value) Early transition: 0.17 ($p = 0.07$) Late transition: 0.17 ($p = 0.25$) Post: 0.66 ($p = 0.56$)
Tom et al (2010) Medical Research Council National Survey of Health and Development	Not reported England, Wales and Scotland	Women born during one week in March 1946 in England, Scotland, and Wales	Social-class stratified, random sample of singleton births during one week in March 1946 Data collected annually for 6 years/7 points observations Baseline $N = 1572$ End of F/U $N = 962$	Composite measure Single item question: "*have you experienced trouble sleeping in the previous 12mths? How much these symptoms bother them*" Responses: "*a lot, a little, not bothered, no trouble sleeping*" 'a lot' and 'little' coded as present	**Demographics** • Educational qualification at the age of 26 • Marital status at age 43 • Parity at the age of 53 **Vasomotor symptoms** • Vasomotor **Physical health** • Somatic symptoms • Night time awakening to use toilet **Psychological symptoms** • Sum score of anxiety, depression, irritability, tearfulness • Life stress **History of trouble sleeping** • Sleep difficulty at the age of 43 **Health behaviour** • Exercise • Smoking • BMI • Alcohol consumption	Generalized estimating equations	Premenopause-premenopause as the reference group ***Moderate sleep difficulty*** Pre-peri: 1.05 (0.79–1.40) Peri-peri: 1.13 (0.86–1.48) Pre/peri-post: 1.27(0.89–1.81) Post-post: 1.13 (0.83–1.54) Pre/peri-HT: 1.20 (0.82–1.77) HT-HT: 0.88 (0.65–1.20) Pre/peri/hyst-hyst: 1.46(1.06–2.02) ***Severe sleep quality*** Pre-peri: 1.40 (0.81–2.43) Peri-peri: 1.97 (1.19–3.28) Pre/peri-post: 2.91 (1.60–5.28) Post-post: 1.88 (1.08–3.27) Pre/peri-HT: 2.64 (1.42–4.90) HT-HT: 1.69 (0.98–2.90) Pre/peri/hyst-hyst: 3.47(1.99–6.04)

Table 1 (*Continued*)

Author & year & study name	Ethnicity & country	Eligibility	Sampling & sample size	Measure of sleep disturbance	Confounders adjusted for	Analysis used	Effect size (95% CI)
Woods and Mitchell (2010) Seattle Midlife Women's Health Study	Multi-ethnicity with the majority being Caucasian USA	• Women responded to ratings of sleep symptoms in their health diaries beginning 1990 • In the late reproductive or menopausal transition • Aged between 35–55 yrs • Not pregnant or lactating • At least one ovary • Understood English	Multistage population based sampling/random selection of household numbers[1] Data collected 8-12 times a year from 1996 to 2000 & quarterly from 2001 to 2005 Only reported 286 for total analysis	Individual sleep problems-DIS, DMS, EMA Questions asking "*the severity of each sleep problem*" Responses: "*0 indicating not present; 4 indicating extreme*" Note: the recall time frame of sleep disturbance is not reported	**Demographics** • Age **Health behaviours** • Smoking • Alcohol • Exercise • Caffeine **Psychological** • Depression • Anxiety • Perceived stress • History of sexual abuse **Physical health** • Perceived health **Hormones** • E2, FSH, Testosterone, epinephrine, norepinephrine **Menopausal symptoms** • Hot flushes **HRT use**	Mixed effects modeling	Premenopause as the reference group Type of effect size: β (*p* value) ***DIS*** Early transition: −0.007 (0.795) Late transition: −0.036 (0.300) Early post: −0.030 (<0.0001) ***DMS*** Early transition: 0.027 (0.420) Late transition: 0.154 (<0.0005) Early post: 0.272 (<0.0001) ***EMA*** Early transition: −0.022 (0.475) Late transition: 0.026 (0.537) Early post: −0.022 (0.689)
Young et al. (2003) Wisconsin Sleep Cohort Study	Caucasian USA	• Aged between 30–60 yrs • Unknown menopausal status	All employees of 5 Wisconsin state agencies (with a full range of job categories from unskilled to professional) Data collected: • Baseline • 4 yrs • 8 yrs Only reported 589 women for total analysis	Individual measurements of sleep problems-DIS, DMs, EMA Questions asking "*how often they experience sleep problems*" Responses: "*never, rarely, sometimes, often, always and almost always*" "often" and "always and almost always" coded as present Note: not recall time frame	**Demographics** • Age • Education **Health behaviours** • BMI • Smoking • Alcohol • Exercise • Caffeine **Physical health** • Health status **Psychological status** • Depression Note: in final model only included age and BMI	**generalized estimating equations** An independence working correlation matrix	Premenopause as the reference group ***DIS*** Peri: 4.18 (1.37–12.77) Post: 2.77 (0.79–9.72) ***DMS*** Peri: 0.59 (0.25–1.43) Post: 1.58 (0.70–3.54) ***EMA*** Peri: 1.43 (0.53–3.89) Post: 0.80 (0.31–2.10) Note: only age and BMI were included in the statistics model
Dennerstein et al. (2007) The Melbourne Women's Midlife Health Project	Caucasian Australia	• Aged between 45–55 yrs • Not on HRT • Had menstrual cycle in the previous 3mths	Population sampling Data collected annually for 9 yrs/10 observations including baseline Only reported 336 women for total analysis	Composite measure Question asks the presence of trouble sleeping, the intensity and frequency in the last 2 weeks Responses: minor, interfering with normal life, debilitating Frequency-the number of days in the last 2 weeks A continuous scale from 0 to 10 was created based on severity and frequency	Could not be determined due to trouble sleeping being both the predicting and outcome variable	Structural equation modelling	A combination of late reproductive and early transition as the reference group Type of effect size: mean (SD) of compound index (frequency and severity) for symptoms Late transition: 0.70 (0.74) Post: 0.71 (0.74) Using HRT: 0.66 (0.74) Note: multiple comparison show significant difference. Confounders adjusted unknown for this analysis

Note: 1. Measure of menopausal transition is presented in Table 3.

Table 2
Assessment of risk of bias.

Author year, & Study name	Selection bias	Baseline response rate	Information bias: menopausal transition	Information bias: sleep	Attrition bias	Selective reporting bias	Adjustment of confounders
Berecki-Gisolf et al. (2009) ALSWH	**Low risk** of bias Women serve as their own control & women were all premenopausal at baseline	**High risk** of bias Response rate: 54% (retrieved from Ref. [3] of their article)	**Low risk** of bias • Consistent with the STRAW criteria • Surgical menopause excluded or censored • HRT use excluded or censored	**High risk** of bias Self-reported sleep difficulty in the previous 12mths	**Low risk** of bias Attrition rate 15%; 1322 of 8612 women recruited lost to follow up	**Low risk** of bias Study website published	Yes See Table 1 for detailed information
Kravitz et al. (2008) SWAN	**Low risk** of bias Women serve as their own control & were all premenopausal or early perimenopausal at baseline	Unclear risk of bias Only mentioned 3302 women were recruited; unclear how many women were invited to participate	**Low risk** of bias • Consistent with the STRAW criteria • Surgical menopausal as a separate group • HRT as a separate group in analysis	**High risk** of bias Self-reported sleep problems in the past 2 weeks	**Low risk** of bias Attrition rate 27.2%; Overall SWAN retention rate was 72.8%	**Low risk** of bias Study protocol available	Yes See Table 1 for detailed information
Owens and Matthews (1998)	**Low risk** of bias Women serve as their own control & were all premenopausal at baseline	Unclear risk of bias Only stated 541 women were recruited, unclear how many were invited to participate	**Low risk** of bias • Consistent with the STRAW criteria • Surgical menopause excluded HRT analysed as a separate group	**High risk** of bias Self-reported sleep problems in the last 6 month	Unclear risk of bias Only stated 213 women for total analysis & unclear the number of women at baseline and F/U	Unclear risk of bias No study protocol or website available	Unclear
Pien et al (2008) Penn Ovarian Aging Study	**Low risk** of bias Women serve as their control & were all premenopausal at baseline	**Low risk** of bias Response rate: 75%; of the 580 eligible women invited, 436 were recruited (Ref 11 of their article)	**High risk** of bias • Not consistent with STRAW • Surgical menopause excluded • Unclear about HRT users	**High risk** of bias Sleep factor score derived from self-reported questionnaire- St. Mary's Hospital Sleep Questionnaire (SMHSQ).	**Low risk** of bias Attrition rate: 26%; 311 of 422 recruited women left at the end of F/U	Unclear risk of bias No study protocol or website available	Yes See Table 1 for detailed information
Tom et al (2010) MRCNSHD	**Low risk** of bias Women serve as their own control & were premenopausal at baseline	**Low risk** of bias Response rate: 87%; Among the 1778 eligible women, 1572 women participated at least once between ages 47–54 (http://www.nshd.mrc.ac.uk/data/response_rates.aspx)	**Low risk** of bias • Consistent with STRAW • Surgical menopause analysed as a separate group • HRT users analysed as a separate group	**High risk** of bias Self-reported sleep difficulty in the past 12mths	Unclear risk of bias Only stated 962 women for total analysis & unclear the number of women lost to follow up	**Low risk** of bias Study website published	Yes See Table 1 for detailed information
Woods and Mitchell (2010) Seattle Midlife Women's Health Study	**Low risk** of bias Women serve as their own control & were premenopausal at baseline	Unclear risk of bias 11222 households were contacted, 508 were recruited & unclear how many were invited (obtained from Seattle Midlife Women Health Study an Overview)	**Low risk** of bias • Consistent with STRAW • Surgical menopause excluded • HRT users analysed as a separate group	**High risk** of bias Self-reported severity of sleep problems	**Low risk** of bias Attrition rate: appx 22%; 286 of 367 recruited women provided data for longitudinal analysis	**Low risk** of bias Study website published	Yes See Table 1 for detailed information
Young et al. (2003) Wisconsin Sleep Cohort Study	**Low risk** of bias Women serve as their own control	**High risk** of bias Response rate: 52%	**High risk** of bias • Consistent with STRAW • Combined surgical with natural HRT users analysed separately	**High risk** of bias Self- reported sleep problems	**Low risk** of bias Attrition rate 20%. Calculated from Response rate being 87% at 4-y F/U & 92% at 8-y F/U	Unclear risk of bias No study protocol or published website	Yes See Table 1 for detailed information
Dennerstein et al. (2007) The Melbourne Women's Midlife Health Project	**Low risk** of bias Women serve as their own control & were at reproductive or early transition stages	**High risk** of bias Response rate: 56%; 779 women invited; 438 women were recruited	**Low risk** of bias • Consistent with the STRAW criteria • Surgical menopause excluded HRT users analysed separately	**High risk** of bias Self- reported sleep index calculated by severity and frequency of sleep problems	**Low risk** of bias Attrition rate: 23.3%; 102 of 438 recruited women lost to follow up	Unclear risk of bias No study protocol or published website	Unclear

Table 3
A comparison of menopausal staging against STRAW criteria.

First author & year	Late reproductive −3a and −3b (premenopausal)	Early menopause −2 (perimenopause)	Late menopause −1 (perimenopause)	Early postmenopause +1a (perimenopause)	Early postmenopause +1b and +1c (postmenopausal)	Late postmenopause +2 (postmenopausal)	Surgical menopause	HRT users
	Regular periods; subtle changes to flow/length of period	Persistent ≥ 7 day change in length between periods	Amenorrhea ≥ 60 days (i.e., 3 months)	Amenorrhea < 12 months	Amenorrhea > 12 months	Amenorrhea > 6 years		
Berecki-Gisolf et al. (2009)	Period in last 3 months and no change in frequency	3–11 months no period or change in frequency			12 months no period		Excluded	Excluded
Kravitz et al. (2008)	Premenopause (menstrual period in the past 3 months and no Decreased predictability)	Early perimenopause: period in the past 3 months but less Predictability in the last 12 months	Late perimenopause (menstrual bleeding in the past 12 months but not in the past 3)		Postmenopause (amenorrhea for the past 12 months)		Analysed as a separate group	Analysed as a separate group
Owens and Matthews (1998)	Regular periods		Missed periods in last 3–11 months		No periods in last 12 months, & no HT		Excluded	Analysed as a separate group
Pien et al (2008)	Premenopause: regular cycles in the normal range (21–35 days) Late premenopause: 1 observed change in cycle length of at least 7 days in either direction compared with the subject's own baseline at enrolment	Early transition: at least 2 cycles with cycle length changes of at least 7 days in cycle length in either direction compared with the subject's own baseline or ≥60 days amenorrhea	Late transition: greater than or equal to 3 months of amenorrhea		Postmenopausal: greater than or equal to 12 months of amenorrhea and no hysterectomy		Excluded	Unknown
Tom et al (2010)	Premenopausal: bleeding in the past 3 months and with the same or increased regularity as in the past year	Perimenopausal: bleeding in the past 12 months but not 3 months or decreased regularity compared to the previous year			Postmenopausal: no reported menstrual bleeding in the past 12 months		Analysed as a separate group	Analysed as a separate group
Woods and Mitchell (2010)	Regular cycles	Persistent irregularity if >6 days	Persistent skipping of 1 or more period	[no definition]	Amenorrhea >12 months and <6 years		Excluded	Analysed as a separate group
Young et al. (2003)	Premenopause, occurrence of period within the previous 3 months	Perimenopause, amenorrhea for at >3 months but <12 months or onset of irregular periods or cycle changes in women who were previously regular			Postmenopause: amenorrhea for ≥12 months or complete hysterectomy or bilateral oophorectomy performed ≥6 months previously		Combined with natural post-menopause	Analysed as a separate group
Dennerstein et al. (2007)	Regular menstrual cycles	Menstruated in the prior 3 months but reported a change in menstrual frequency	At least 3 months of amenorrhea but less than 12 months of amenorrhea		Amenorrhea for more than 12 months		Excluded	Analysed as a separate group

Databases and reference lists searched; 1480 articles located.
Studies excluded (*N* = 71).
Abstracts excluded (*N* = 198).
Studies excluded (*N* = 2).

other four studies were rated as having unclear risk due to the availability of protocol or published websites [10,12,13,15].

4. Discussion

Based on our inclusion criteria, eight studies provided data in evaluating the longitudinal relationship between menopausal transition and sleep disturbance. Five studies measured sleep disturbance as a global construct. Of these five studies, three found a significant but small increase in the odds of experiencing sleep disturbance from premenopause to perimenopause [9,10,16], and four reported similar results for the transition to postmenopause [9,10,12,16]. The other three studies measured common sleep problems including DIS, DMS and EMA. Of these three studies, two consistently found a significant increase in the likelihood of having DMS when women transition from premenopause to both perimenopause and postmenopause. Findings on DIS and EMA were inconsistent, which makes it difficult to draw conclusions due to the small number of studies. Results from the current review of longitudinal studies verified the findings in our previous review of the cross-sectional relationship between menopausal stages and sleep disturbance; that is, perimenopause and postmenopause are associated with worse sleep compared to premenopause. Furthermore, a causative relationship between menopausal transition and sleep disturbance is indicated due to the longitudinal nature of the reviewed data and the statistical control of identified confounders. Unfortunately, we could not analyse the influence of cross-cultural variation, as there was little corresponding information on non-Caucasian women with which to compare.

The five studies that evaluated sleep disturbance as a global construct were similar in the majority of study characteristics except for the measurement of sleep disturbance. Four of these five studies used a single item question on sleep disturbance to measure sleep disturbance and found significant outcomes; whereas the Penn Ovarian Aging Study used sleep quality score as the measurement and yielded non-significant outcomes [13]. Sleep quality score was a component score generated from the factor analysis of the St Mary's Hospital Sleep Questionnaire. Although it accounted for 37% of the variance of the entire questionnaire, its correlation with a single item question on sleep disturbance was moderate ($r = 0.44$ to 0.50), indicating the underlying structures measured by this sleep quality factor score and a single item on sleep disturbance might not be the same. In fact, the subgroup analysis in our previous review on the cross-sectional data has constantly found a smaller OR for validated sleep measure than non-validated ones [7].

Of the three studies that examined the relationships between individual sleep problems and menopausal transition, the Wisconsin Sleep Cohort Study [15] reported by large non-significant findings in contrast to the SWAN [11] and the Seattle Midlife Women's Health Study [14]. Careful comparison of these studies revealed considerable difference in two aspects. First, there were only two follow up points in the Wisconsin Sleep Cohort Study over an eight-year period of time. This is understandable from a feasibility point of view, as women in this study were required to be assessed in the sleep lab to evaluate objective sleep disturbance in relation to menopausal transition [15]. Nevertheless, the number of follow up time points was much fewer compared to the eight or more follow up periods for other studies over a similar duration. The small number of observations combined with the relatively small sample size in this study increases the likelihood of Type II error, i.e., increasing the likelihood of a false non-significant outcome. Second, in the statistical analysis of the Wisconsin Sleep Cohort Study, the authors selected the independence working correlation structure as the correlation of covariance in analysis, whereas the other included studies set the working correlation structure as exchangeable or equi-correlational. An independence working correlation structure assumes independence between observations and overestimates the standard error of time-dependent variable [17], which, in this case, is menopausal stage. Although generalised estimating equations analysis may be robust to violations of the assumption of independence [17], again, along with fewer observations, this may increase the likelihood of Type II error.

Despite the small number of longitudinal studies available, the studies included in this review were robust. The risk of selection bias was low in all of the included studies due to the nature of longitudinal design. Risk of information bias was high for sleep disturbance, because the measurement was based on self-reported responses. However, subjective sleep disturbance is only able to be reported by participants themselves. The studies included in this review have mostly adhered to the STRAW classification of menopausal stages, except for the Penn Ovarian Aging Study [13], where there was no clear differentiation between early and late transition according to their definition. That the Wisconsin Sleep Cohort Study combined natural and surgical menopause into one group is concerning, as literature has clearly shown that women with surgical menopause often experience more frequent and severe menopausal symptoms [7,18]. In addition, their previous health status differs significantly from women with natural menopause.

Attrition rates overall were lower than 30%. The majority of the included studies controlled for extensive range of confounders. In the study undertaken by Owens and Matthews, no confounders were controlled for, and this was perhaps reflected in a relatively greater OR of 2.33 [12].

5. Conclusions

Based on the available literature, there is reasonable evidence that there is an increased risk of sleep disturbance among midlife women as they go through natural menopausal transition, beyond the effect of age, health behaviour (e.g., physical activity, alcohol consumption, and tobacco use), physical health (e.g., number of health conditions, number of health symptoms), psychological health (e.g., depressive symptoms, anxious symptoms), and vasomotor symptoms. There may also be an increased risk of sleep disturbance among women as they transition through natural menopausal, beyond the effect of previous sleep disturbance [16], estrodiol and FSH [11]; although these conclusions are less strong as they are based on a single study alone. Although menopause is a clear risk factor for sleep disturbance in women, nonetheless, the magnitude attributed to the menopausal transition is small. In the current review, the studies included were of mostly of rigorous design, which supports the conclusion that menopause is a causal factor in sleep disturbance for women.

Contributors

Qunyan Xu is responsible for the:

1. Conception of the systematic review.
2. Design of the systematic review.
3. Conduction of literature search and article screening.
4. Description of included studies in terms of important characteristics.
5. Critique of the methodological quality of included studies.
6. Drafting parts of the article.
7. Revision of the article.

Cathryne Lang is responsible for the:

1. Conduction of literature search and article screening.
2. Description of included studies in terms of important characteristics.

3. Critique of the methodological quality of included studies.
4. Drafting parts of the article.
5. Revision of the article.

Naomi Rooney is responsible for the:

1. Description of included studies in terms of important characteristics.
2. Drafting parts of the article.
3. Revision of the article.

Competing interest

The authors declare no conflict of interest.

Provenance and peer review

Not commissioned; externally peer reviewed.

Funding

This research is funded by 2013 Faculty of Arts and Science, Early Career Research Incentive Award, Australian Catholic University.

Acknowledgements

We are grateful for the faculty grant from the Australian Catholic University which facilitated this research.

References

[1] Phillips BA, Collop NA, Drake C, Consens F, Vgontzas AN, Weaver TE. Sleep disorders and medical conditions in women. J Womens Health 2008;17:1191–9.
[2] Lindberg E, Janson C, Gislason T, Bjornsson E, Hetta J, Boman G. Sleep disturbances in a young adult population: can gender differences be explained by differences in psychological status. Sleep 1997;20:381–7.
[3] Polo-Kantola P. Sleep problems in midlife and beyond. Maturitas 2011;68:224–32.
[4] Guidozzi F. Sleep and sleep disorders in menopausal women. Climacteric 2013;16:214–9.
[5] Ameratunga D, Goldin J, Hickey M. Sleep disturbance in menopause. Intern Med J 2012;42:742–7.
[6] Joffe H, Massler A, Sharkey KM. Evaluation and management of sleep disturbance during the menopause transition. Semin Reprod Med 2010;28:404–21.
[7] Xu Q, Lang CP. Examining the relationship between subjective sleep disturbance and menopause: a systematic review and meta-analysis. Menopause; 2014.
[8] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.
[9] Berecki-Gisolf J, Begum N, Dobson AJ. Symptoms reported by women in midlife: menopausal transition or aging. Menopause 2009;16:1021–9.
[10] Dennerstein L, Lehert P, Guthrie JR, Burger HG. Modeling women's health during the menopausal transition: a longitudinal analysis. Menopause 2007;14:53–62.
[11] Kravitz HM, Zhao X, Bromberger JT, Gold EB, Hall MH, Matthews KA, et al. Sleep disturbance during the menopausal transition in a multi-ethnic community sample of women. Sleep 2008;31:979–90.
[12] Owens JF, Matthews KA. Sleep disturbance in healthy middle-aged women. Maturitas 1998;30:41–50.
[13] Pien GW, Sammel MD, Freeman EW, Lin H, DeBlasis TL. Predictors of sleep quality in women in the menopausal transition. Sleep 2008;31:991–9.
[14] Woods NF, Mitchell ES. Sleep symptoms during the menopausal transition and early postmenopause: observations from the Seattle Midlife Women's Health Study. Sleep 2010;33:539–49.
[15] Young T, Rabago D, Zgierska A, Austin D, Laurel F. Objective and subjective sleep quality in premenopausal, perimenopausal, and postmenopausal women in the Wisconsin Sleep Cohort Study. Sleep 2003;26:667–72.
[16] Tom SE, Kuh D, Guralnik JM, Mishra GD. Self-reported sleep difficulty during the menopausal transition: results from a prospective cohort study. Menopause 2010;17:1128–35.
[17] Twisk JWR. Applied longitudinal data analysis for epidemiology: a practical guide. New York, NY: Cambridge University Press; 2003.
[18] Bhattacharya SM, Jha A. A comparison of health-related quality of life (HRQOL) after natural and surgical menopause. Maturitas 2010;66:431–4.